Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
J. Math. Anal. Appl. 303 (2005) 622–626
Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS
www.elsevier.com/locate/jmaa

A new proof of the strong duality theorem for semidefinite programming ☆

Qingzhi Yang

School of Mathematics and LPMC, Nankai University, Tianjin 300071, PR China

Received 1 July 2003
Available online 19 November 2004
Submitted by M.A. Noor

Abstract

Semidefinite programs are convex optimization problems arising in a wide variety of applications and are the extension of linear programming. Most methods for linear programming have been generalized to semidefinite programs. Just as in linear programming, duality theorem plays a basic and an important role in theory as well as in algorithmics. Based on the discretization method and convergence property, this paper proposes a new proof of the strong duality theorem for semidefinite programming, which is different from other common proofs and is more simple.

Keywords: Semidefinite programming; Duality; Semiinfinite programming; Discretization

Consider the following *semidefinite programming problem* (SDP):

$$\text{(P)} \quad \min c^T x \quad \text{s.t. } F(x) \succcurlyeq 0, \tag{0.1}$$

where $F(x) := F_0 + \sum_{i=1}^m x_i F_i$, $c \in R^m$, $F_i \in R^{n\times n}$, $i = 0, 1, \ldots, m$, are symmetric and F_i, $i = 1, \ldots, m$, linearly independent. By "$\succcurlyeq$" we denote the Löwner partial order, i.e., $F(x) \succcurlyeq 0$ means that $y^T F(x) y \geqslant 0$ for all $y \in R^n$. Obviously, the feasible set of (P) is convex.

☆ This work is supported by the National Natural Science Foundation of China, Grant 10171055.
E-mail address: qz-yang@nankai.edu.cn.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2004.08.063

Semidefinite programs are an important class of convex optimization which directly arises in a number of important applications such as control theory, combinatorial optimization, matrix theory, etc. Many convex optimization problems, e.g., linear programming and (convex) quadratically constrained quadratic programming can be cast as semidefinite programs [1,3,8]. As in linear programming, dual theorem plays a basic and an important role in theory as well as in algorithmics. In this note we intend to provide a new proof of dual theorem for semidefinite programming based on discretization of semiinfinite programming [4–7], though this result has been known for a long time [1,3].

The dual problem of (P) is the following [1,3]:

$$\text{(D)} \quad \max\ -\operatorname{Tr} F_0 Z \quad \text{s.t. } \operatorname{Tr} F_i Z = c_i,\ Z \succcurlyeq 0, \tag{0.2}$$

where Tr denotes the trace of a matrix.

Obviously, the equivalent linear semiinfinite programming (LSIP) form of (P) is

$$(\mathrm{P}') \quad \min c^T x \quad \text{s.t. } a^T(y)x + b(y) \geqslant 0,\ \text{for any } y \in Y, \tag{0.3}$$

where $Y = \{y \in R^n \mid y^T y = 1\}$, $a^T(y) = (y^T F_1 y, y^T F_2 y, \ldots, y^T F_m y)$, $b(y) = y^T F_0 y$.

Roughly speaking, in the discretization method we choose a finite grid set $Y_d \subset Y$ to replace the infinite index set Y in semiinfinite programming (SIP) [2,5], and get the discretized problem of (SIP). Thus one can use constrained optimization methods to solve the discretized problem in order to obtain an approximate solution of (SIP) for sufficient small d, where d denotes the Hausdorff distance between Y and Y_d, i.e.,

$$d := \operatorname{dist}(Y_d, Y) = \max_{y \in Y} \min_{\hat{y} \in Y_d} \|\hat{y} - y\|.$$

It measures the fineness of the mesh grid Y_d. Discretization provides a way to solve (SIP) [2,5]. However in this note we want to propose a more simple proof of the strong duality theorem for (SDP) with the idea of discretization.

The discretized problem of (P′) is

$$(\mathrm{P}_d) \quad \min c^T x \quad \text{s.t. } a^T(y)x + b(y) \geqslant 0,\ \text{for any } y \in Y_d, \tag{0.4}$$

where $Y_d \subset Y := \{y \in R^n \mid y^T y = 1\}$ is a grid.

First we recall the following denotations and definitions [1–3]:

$$X^P(Y) = \bigl\{x \mid y^T F(x) y \geqslant 0,\ \text{for any } y \in Y\bigr\} \subset R^n.$$

It can be seen that $X^P(Y)$ is convex. Let $Y_d \subset Y$, $|Y_d| < \infty$,

$$X^P(Y_d) = \bigl\{x \mid y^T F(x) y \geqslant 0,\ \text{for any } y \in Y_d\bigr\}.$$

Definition 1. P-level sets (for level κ) are defined by

$$L_{\succcurlyeq}(X^P, c, \kappa) = \bigl\{x \in X^P(Y) \mid c^T x \leqslant \kappa\bigr\}.$$

Definition 2. The problem (P) is strictly feasible if there exists an x with $F(x) \succ 0$. Here $F(x) \succ 0$ means that the matrix is positive definite.

Definition 3. The problem (D) is strictly feasible if there exists a Z with $Z = Z^T \succ 0$, $\operatorname{Tr} F_i Z = c_i$, $i = 1, \ldots, m$.

Lemma 1. *If* (P) *is feasible and* (D) *is strictly feasible, then all the P-level sets are bounded.*

Proof. Since (P) is strictly feasible, there is an $\tilde{x}$ with

$$y^T F(\tilde{x}) y > 0, \quad \text{for all } y \in Y.$$

If the statement is false, we must have a P-level set denoted by

$$L_{\succcurlyeq}(X^P, c, \kappa) = \{x \in X^P \mid c^T x \leqslant \kappa\}$$

which is unbounded. Then there exists a sequence of points $x^k \in L_{\succcurlyeq}(X^P, c, \kappa)$ such that $\|x^k\| \to \infty$ for $k \to \infty$. We can write x^k in the form

$$x^k = x^1 + t_k h^k, \quad \text{with } \|h^k\| = 1,\ t_k > 0,\ t_k \to \infty.$$

By choosing a subsequence (if necessary) we can assume $h^k \to h$, $\|h\| = 1$ and write $h^k = h + v^k$, $v^k \to 0$. By feasibility of x^k it follows that

$$F_0 + \sum_{i=1}^m x_i^1 F_i + t_k \left(\sum_{i=1}^m \left(h_i + v_i^k\right) F_i \right) \succcurlyeq 0.$$

After division by t_k and letting $k \to \infty$ we deduce $\sum_{i=1}^m h_i F_i \succcurlyeq 0$. With the condition that (D) is strictly feasible, we have some $Z = Z^T \succ 0$ such that $\operatorname{Tr} F_i Z = c_i$, $i = 1, \dots, m$. So we have

$$c^T h = \sum_{i=1}^m h_i \operatorname{Tr}(F_i Z) = \operatorname{Tr}\left(\sum_{i=1}^m h_i F_i \right) Z \geqslant 0.$$

If $c^T h = 0$, then by the property of trace [1] we have that $(\sum_{i=1}^m h_i F_i) Z = 0$. Since $(\sum_{i=1}^m h_i F_i) \succcurlyeq 0$, $Z \succ 0$, we can conclude that $\sum_{i=1}^m h_i F_i = 0$. It contradicts the fact that F_i, $i = 1, \dots, m$, are linearly independent. If $c^T h > 0$, then $c^T x^k = c^T x^1 + t_k (c^T h + c^T v^k) \to \infty$ as $k \to +\infty$. It contradicts the fact that

$$\{x + th\}_{t > 0} \subset L_{\succcurlyeq}(X^P, c, \kappa) = \{x \in X^P \mid c^T x \leqslant \kappa\}.$$

Hence the statement is true. □

Lemma 2. *If* (P) *is feasible and* (D) *is strictly feasible, then for every $\epsilon > 0$, there exists $\delta_\epsilon > 0$ such that for $Y_d \subset Y$, $d(Y_d) < \delta_\epsilon$, it is the case that* (P_d) *is solvable and for every solution x_d, there exists a solution x^* of* (P) *such that $\|x_d - x^*\| < \epsilon$.*

Proof. Notice that the problem (P) is equivalent to the problem (P′), and with Lemma 1 and Theorem 4.4 in [2], the result follows. □

Remark. The dual (D′) of (P′) can be got easily, and by Lemma 12.6 in [3] we know that (D) is strictly feasible iff $c \in \operatorname{intcone}\{a(y) \mid y \in Y\}$ in the notation of [2] or [3]. So (D′) is superconsistent [2]. By Theorem 6.11 in [2] if (P) is consistent it follows that (D) is strictly feasible iff P-level sets are bounded.

Let val(P) be the optimal value of the problem (P), val(D) be the optimal value of the problem (D). Then one has the following strong duality theorem.

Theorem 1. *If* (P) *is feasible and* (D) *is strictly feasible, then* $\mathrm{val(P)} = \mathrm{val(D)}$. *Moreover a solution* x^* *of* (P) *exists.*

Proof. By Lemma 1 the P-level sets are bounded and closed thus compact (and nonempty for κ large enough). Consequently, a solution x^* of (P) exists.

Denote $Y_d = \{y_1, y_2, \ldots, y_l\}$, then (P_d) is a linear optimization problem with $l = l_d$ constraints, i.e.,

$$\min c^T x \quad \text{s.t. } Ax \leqslant b,$$

where

$$A = -\begin{pmatrix} y_1^T F_1 y_1 & \cdots & y_1^T F_m y_1 \\ \vdots & \ddots & \vdots \\ y_l^T F_1 y_l & \cdots & y_l^T F_m y_l \end{pmatrix}, \qquad b = \begin{pmatrix} y_1^T F_0 y_1 \\ \vdots \\ y_l^T F_0 y_l \end{pmatrix}. \tag{0.5}$$

With duality of linear programming, we get the dual of (P_d):

$$(\mathrm{D}_d) \quad \max\ -\mathrm{Tr}\, F_0 Z \quad \text{s.t } \mathrm{Tr}\, F_i Z = c_i,\ i = 1, \ldots, m,\ Z \in Z_d, \tag{0.6}$$

where $Z_d = \{Z \mid Z = \sum_{j=1}^{l} t_j y_j y_j^T,\ t_j \geqslant 0,\ j = 1, \ldots, l\}$.

Let $\mathrm{val}(\mathrm{P}_d)$ be the value of the problem (P_d), and $\mathrm{val}(\mathrm{D}_d)$ be the value of the problem (D_d). For every feasible x and Z to (P) and (D), respectively, we have

$$c^T x - (-\mathrm{Tr}\, F_0 Z) = \mathrm{Tr}\left(F_0 + \sum_{i=1}^{m} x_i F_i\right) Z \geqslant 0.$$

Therefore,

$$\mathrm{val(P)} \geqslant \mathrm{val(D)}. \tag{0.7}$$

Let x^* be an optimal solution of (P). By Lemma 2, we can generate solutions x_d of problem (P_d) such that $x_d \to x^*$ as $d \to 0$. Since the feasible set of (P) is included in the feasible set of (P_d), it follows that

$$\mathrm{val}(\mathrm{P}_d) \leqslant \mathrm{val(P)}. \tag{0.8}$$

At the same time, by using the Lemma 2 we have $\mathrm{val}(\mathrm{P}_d) \to \mathrm{val(P)}$ as $d \to 0$. So $\mathrm{val}(\mathrm{P}_d)$ is bounded for sufficient small d. Now with dual theorem of linear programming we have $\mathrm{val}(\mathrm{P}_d) = \mathrm{val}(\mathrm{D}_d)$. While the feasible set of (D_d) is included in the feasible set of (D), so

$$\mathrm{val}(\mathrm{D}_d) \leqslant \mathrm{val(D)}. \tag{0.9}$$

Letting d tend to zero yields

$$\mathrm{val(P)} = \lim \mathrm{val}(\mathrm{P}_d) = \lim \mathrm{val}(\mathrm{D}_d) \leqslant \mathrm{val(D)}. \tag{0.10}$$

Therefore $\mathrm{val(P)} = \mathrm{val(D)}$ from (0.7). $\square$

Acknowledgment

I thank the referee for helpful suggestions.

References

[1] L. Vanderberghe, S. Boyd, Semidefinite programming, SIAM Rev. 38 (1995) 49–95.
[2] R. Hettich, K. Kortanek, Semi-infinite programming: Theory, methods, and application, SIAM Rev. 35 (1994) 380–429.
[3] U. Faigle, W. Kern, G. Still, Algorithm Principles of Mathematical Programming, Kluwer Academic, Dordrecht, 2002.
[4] L. Vanderberghe, S. Boyd, Connection between semi-infinite and semidefinite programming, in: R. Reemtsen, J.-J. Rückmann (Eds.), Semi-Infinite Programming, Kluwer Academic, Dordrecht, 1998, pp. 277–292.
[5] G. Still, Discretization in semi-infinite programming: the rate of convergence, Math. Programming Ser A. 91 (2001) 53–68.
[6] R. Hettich, An implementation of a discretization method for semi-infinite programming, Math. Programming 34 (1986) 354–361.
[7] R. Hettich, G. Gramlich, A note on an implementation of a method for quadratic semi-infinite programming, Math. Programming 46 (1990) 249–254.
[8] F. Alizadeh, Interior point methods in semidefinite programming with applications to combinatorial optimization, SIAM. J. Optim. 5 (1995) 13–51.